CottonGen
Cotton Database Resources

January 2023

Welcome to the 1st issue of the CottonGen newsletter in 2023. This newsletter is issued to inform users about **new or updated data and tools in CottonGen**. In addition to new and updated data, each issue will provide more information on data or tools on **the featured tools/data** section.

New Genome Data/Functional Analysis

Data from **whole genome assemblies**

G. arboreum (A2) genome HAU v2 (*Wang et al., 2022*)

CottonGen Functional Analysis (InterProScan and KEGG, Protein Homologies, Cottongen Marker and RefTran Alignments, etc.) **added to** the **genome:**

G. herbaceum (A1) 'Wagad' genome USDA_v1 *(Ramaraj et al., 2022)*

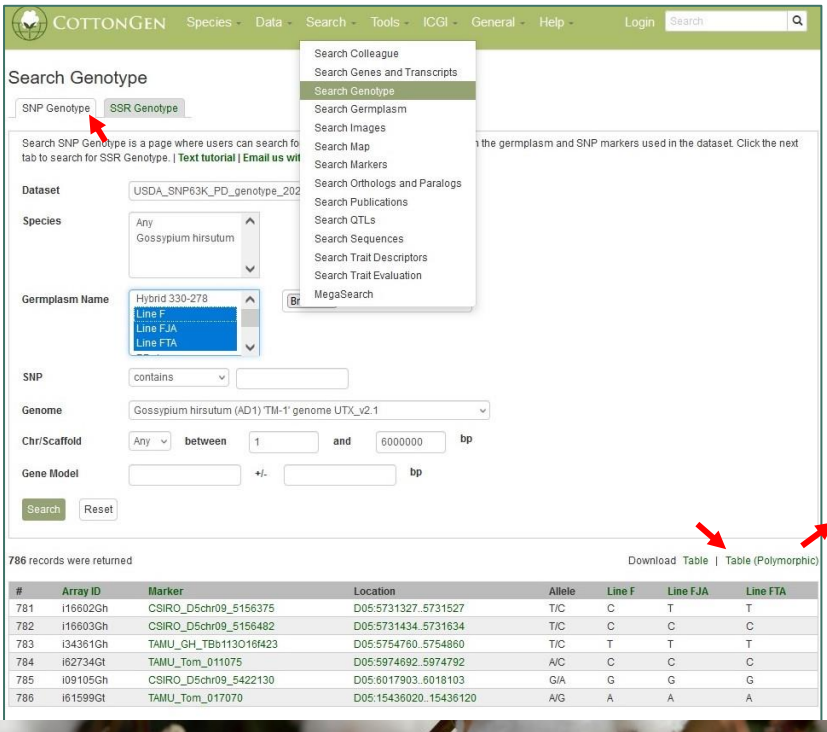
G. anomalum (B1) genome JAAS_v1.2_a1.1 (*Xu 2022*)

G. stocksii (E1) genome_ZSTU v1. *(Yu et al.,2021)*

G. bickii (G1) genome JZU_v1. *(Sheng et al., 2022)*

Other New Data

- Over 6.4K RBTN non-fiber trait data for 13 years from 10 difference locations (RBTN).
- Over 25K SNP/InDel markers and genome positions
- Over 1700 new germplasm or existing germplasm have new assigned GRIN_IDs or the publications from new germplasm registrations.
- Updated US National and Statewide Cotton Variety Test page
- Updated Species Summary page and links with organized names of genome and genome group



Genotype Search

In the interface of Search|Genotype|SNP, select a Dataset name from the dropdown list, a list of germplasm used in this dataset is given. Choose one or more germplasm and click 'Search', the result is given (below-left). An option of 'Table (Polymorphic)' allows user to view/download the table of polymorphisms-only among the selected germplasm.

#	Array ID	Marker	Location	Allele	Line F	Line FJA	Line FTA
1	i30433Gh	TAMU_GH_TBb075C15f202	D05:27287751..27287851	A/G	G	A	A
2	i23619Gh	TAMU_GH_TBb006O03f688	D05:293930..294030	A/C	C	C	A
3	i08789Gh	CSIRO_D5chr09_436357	D05:473417..473617	T/G	T	T	G
4	i21841Gh	USDA_CFB4573	D05:1962621..1962708	G/A	G	G	A
5	i08857Gh	CSIRO_D5chr09_1986488	D05:2227234..2227434	C/T	C	C	T
6	i08914Gh	CSIRO_D5chr09_2955699	D05:3379515..3379715	T/C	C	T	T
7	i16546Gh	CSIRO_D5chr09_2456603	D05:2868015..2868215	A/G	A	A	G
8	i08887Gh	CSIRO_D5chr09_2456748	D05:2868160..2868360	C/T	C	C	T
9	i61131Gt	TAMU_Tom_011056	D05:2868954..2869054	T/C	T	T	C
10	i08888Gh	CSIRO_D5chr09_2460076	D05:2871488..2871688	G/A	A	A	G
11	i42437Gh	TAMU_GH_TBh073P08f192	D05:3231345..3231445	T/C	T	C	C
12	i36700Gh	TAMU_GH_TBh018F13f110	D05:2880924..2881024	A/C	C	C	A
13	i08911Gh	CSIRO_D5chr09_2819127	D05:3257939..3258139	A/G	G	A	G
14	i08920Gh	CSIRO_D5chr09_3005887	D05:3431321..3431521	C/A	A	C	C
15	i08919Gh	CSIRO_D5chr09_3003221	D05:3428661..3428861	T/C	C	T	T
16	i08969Gh	CSIRO_D5chr09_3700792	D05:4167178..4167378	A/G	A	G	A
17	i08965Gh	CSIRO_D5chr09_3689945	D05:4063921..4064121	A/G	G	A	G
18	i16569Gh	CSIRO_D5chr09_3690110	D05:4155475..4155675	A/G	A	G	A
19	i08992Gh	CSIRO_D5chr09_4060086	D05:4389436..4389636	G/A	A	G	G
20	i47407Gh	TAMU_GH_TBh119P20r314	D05:4528972..4529072	T/C	T	C	C
21	i08996Gh	CSIRO_D5chr09_4066862	D05:4525720..4525920	T/G	T	G	G
22	i45318Gh	TAMU_GH_TBh100L12f359	D05:4487384..4487484	A/C	C	A	C
23	i09022Gh	CSIRO_D5chr09_4382170	D05:4926991..4927191	C/A	C	A	C
24	i09024Gh	CSIRO_D5chr09_4399447	D05:4941675..4941875	A/G	G	A	G
25	i14148Gh	CSIRO_D5Chr09_4402089	D05:4944331..4944531	G/A	G	A	G
26	i08786Gh	CSIRO_D5chr09_256427	D05:259900..260100	C/T	C	C	T
27	i51819Gb	TAMU_Gb379_010910	D05:340510..340608	T/G	K	T	T
28	i08796Gh	CSIRO_D5chr09_556083	D05:601551..601751	G/A	A	A	G

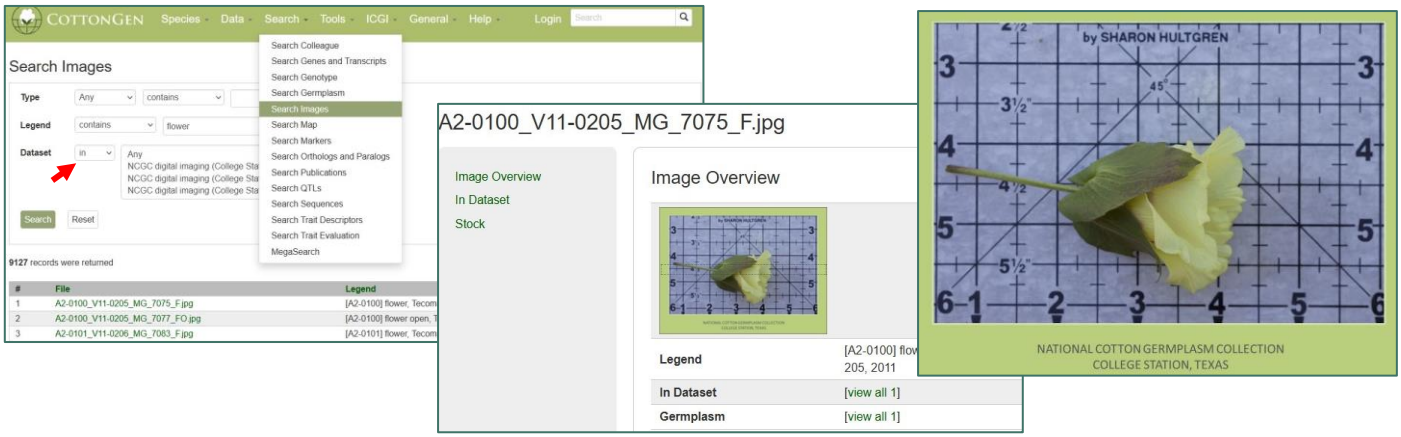
CottonGen
Cotton Database Resources

January 2023

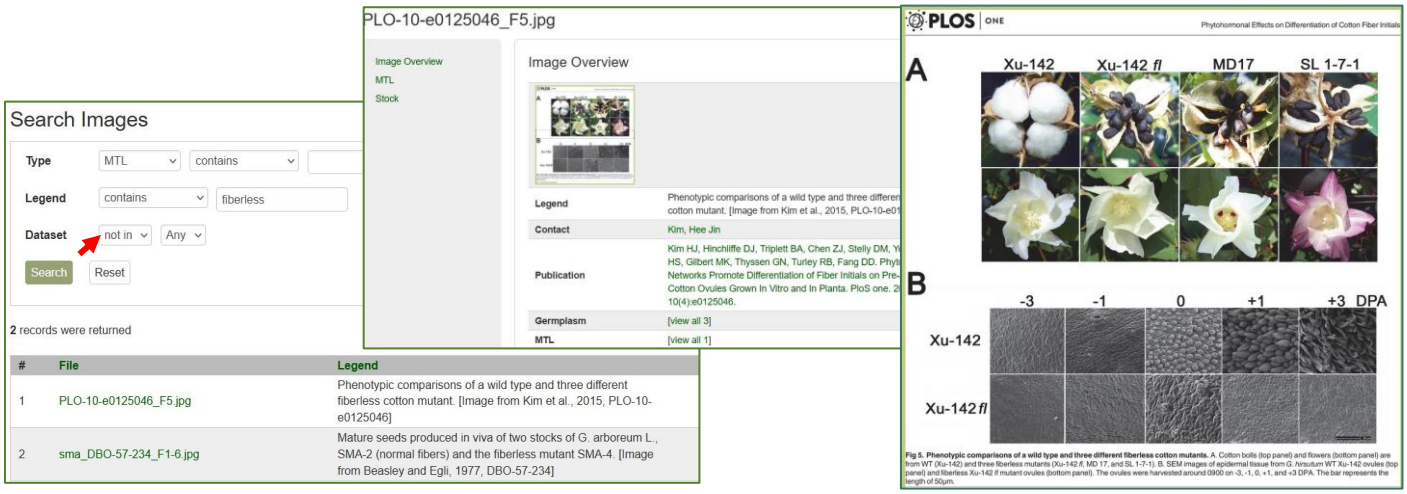
Image Searches

CottonGen currently contains 45,211 digital images. 45,198 of them are from USDA ARS National Cotton Germplasm Characterization (NCGC), while others are from different sources based on different purposes, such as to describe complex pedigrees, to describe mutant or disease symptoms, etc. For the convenience of users, in the image Search we use all NCGC images belong to 'dataset' (i.e. NCGC) and the remaining images do not belong to any 'dataset'. Here are examples of the Search Images functionality:

EXP1. Find NCGC germplasm flower images: Search Images->Legend contains 'flower' -> Dataset 'In' (in NCGC image datasets).



EXP2. Find fiberless mutant images: Search Images->Type=MTL-> Legend contains 'fiberless' ->Dataset 'not in' (i.e.,not NCGC)



Join the [CottonGen Mailing List](#) and follow us on [Twitter](#)

Funded by:
Cotton Inc.; USDA-ARS; Bayer CS; CORTEVA; USDA National Research Project (NRSP10)